



Is HighSpeed TCP a way?

Evandro de Souza

NERSC/DSD

August 2002

Supervisor: Deb Agarwal



Outline

Motivation

TCP Congestion Control

HighSpeed TCP Proposal

Research and Results

Conclusion



DOE Vision

The vision for DOE scientific networking is that major DOE applications and facility-based experiments will interconnect with widely distributed terascale supercomputing, petascale storage, high performance visualization, and remote collaborators, to dynamically create virtual laboratories.



Expect Performance

Most effective and the highest possible performance networking where needed by its high impact science

Expect performance related to the present Esnet

- Bandwidth 4 to 16 times greater
- Links of 2.5 to 10 Gbps



Network Issues

Allow high performance end-to-end, at speeds from 100 Mbps now to 1 Gbps and above for individual flows

New developments of resilient network protocols that match these speeds

New concepts of fairness and new modes of network operation, monitoring and management



Congestion Control History

Transmission Control Protocol (TCP) is the most commonly used transport protocol in the Internet

In the first years it had only a very rudimentary congestion control mechanism

The Internet experienced several severe congestion collapses in the mid 80s



Congestion Collapses

The cause for the congestion collapses was examined by Von Jacobson in 88, and proposed a modified congestion control mechanism

TCP is still in use today because its robust congestion control mechanism

This mechanism causes TCP to reduce its sending rate when congestion is encountered along the network path, as evidenced by dropped packets.



TCP Congestion Control

With recent developments in high-speed networking and applications that use TCP, performance issues in TCP are of increasing interest and importance

The congestion avoidance mechanism of TCP adopts a window-based flow control, which controls the number of on-the-fly packets in the network

The source terminal is allowed to send the number of packets given by its window size (cwnd)



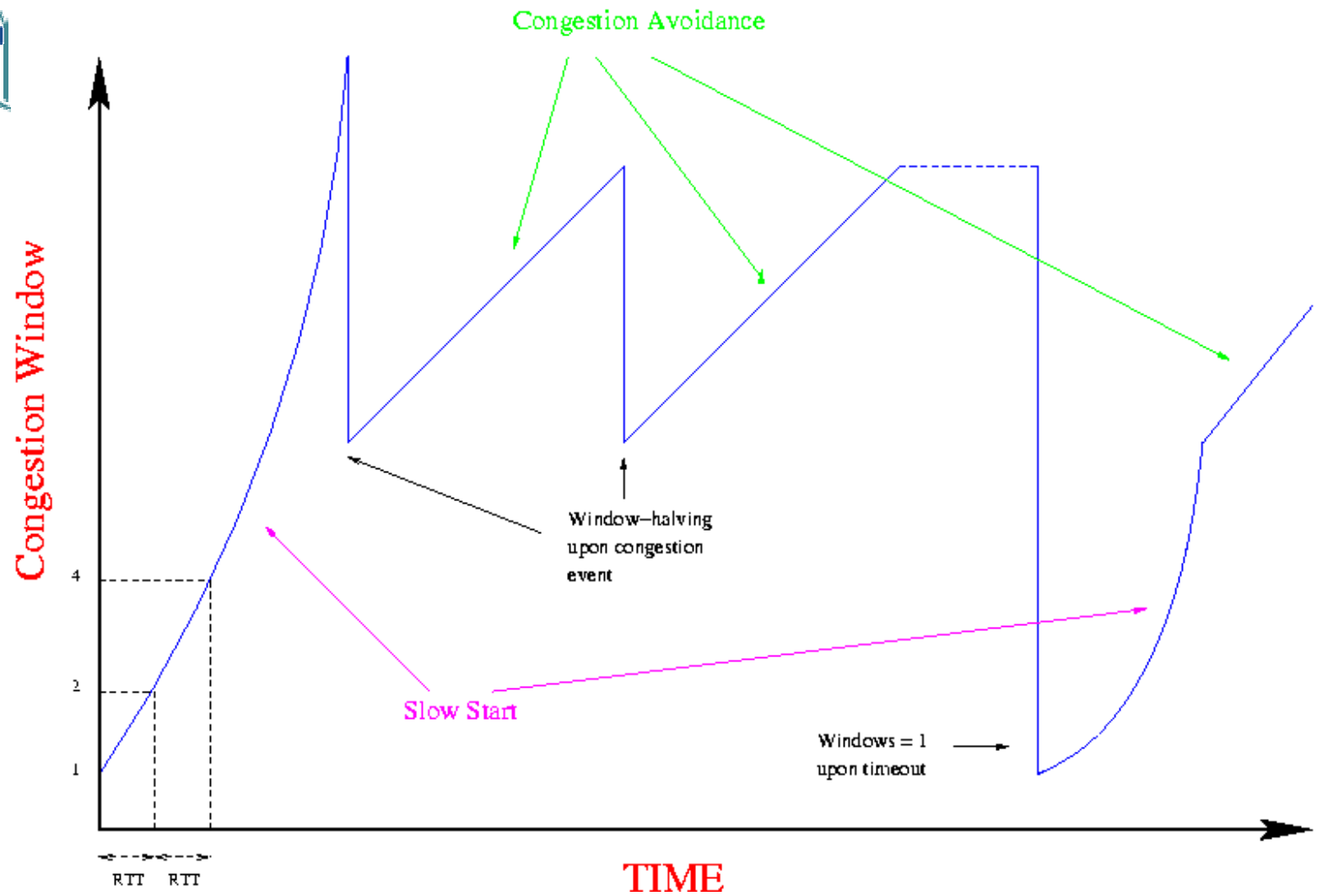
TCP Congestion Control

The window size is updated at the receipt of ACK (ACKnowledgement) packet

The key idea of the congestion avoidance mechanism of TCP is to dynamically control the window size according to severity of the congestion in the network



TCP CONGESTION CONTROL





TCP Response Function

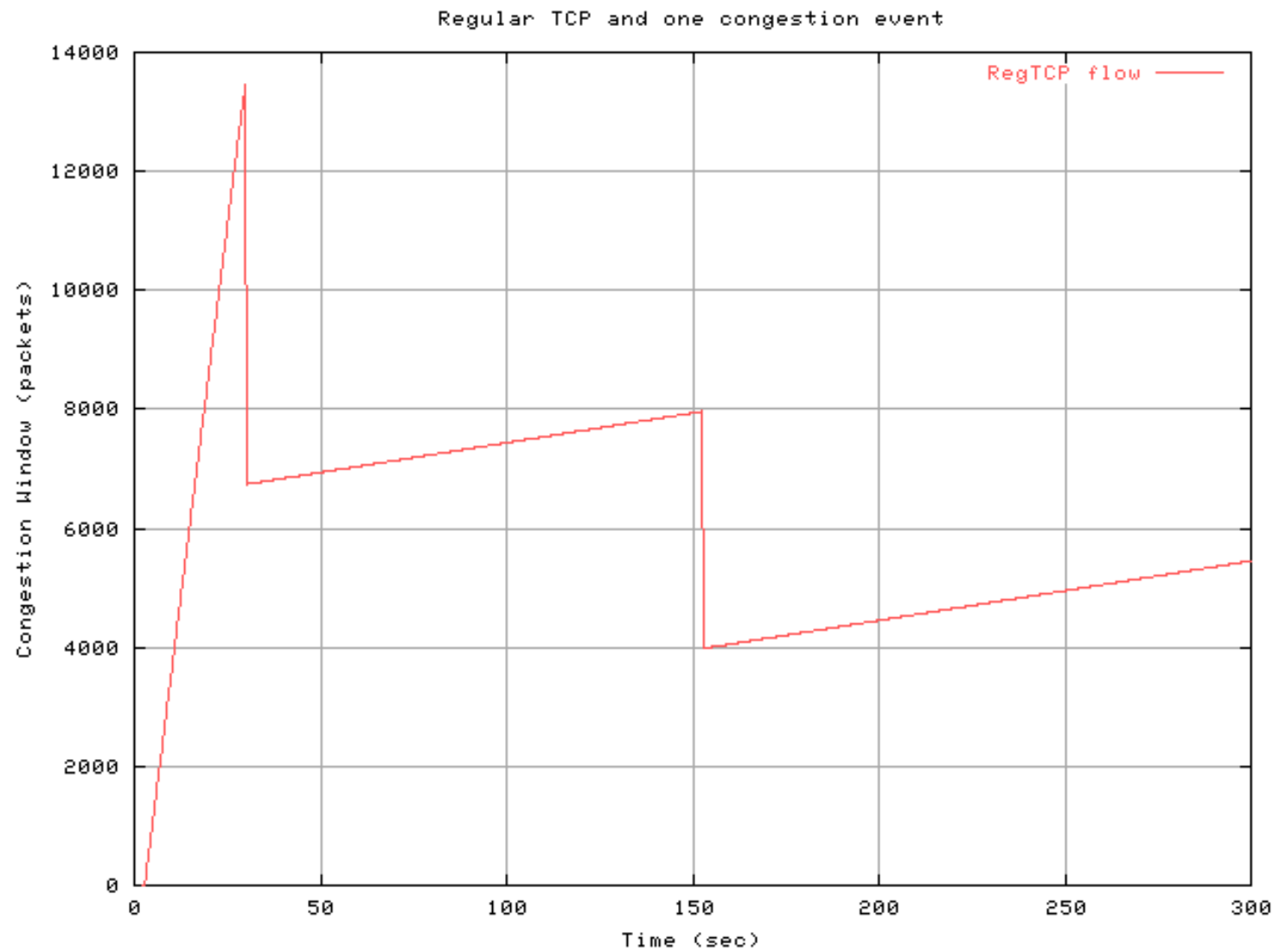
It has been known that the current TCP does not scale to high-speed networks directly
TCP's response function:

$$S = \frac{\sqrt{1.5}}{\sqrt{p}} \text{ pkts} / RTT$$



High Speed Problem

For example, for a TCP connection with 1500-byte packets and a 100 ms round-trip time, filling a 10 Gbps pipe would require a congestion window of 83,333 packets, and a packet drop rate of at most one drop every 5,000,000,000 packets. This is at most one drop per $S = 6000s$, or 1h:40m





TCP-Friendliness

It is easy to design congestion control schemes that achieve higher sending rates at a given loss rate. However, the challenge is to do so while retaining the TCP-compatibility (or TCP-friendliness) properties of the congestion control algorithm



HSTCP Proposal

A modification to TCP's congestion control mechanism for use with TCP connections with large congestion windows. [Sally Floyd/ICIR]

It aims to achieve high per-connection throughput without requiring unrealistically low packet loss rates and TCP-compatible performance in environments with moderate or high congestion



AIMD Algorithm

Basic algorithm to adjust the TCP congestion window (Von Jacobson 88)

- ACK: $w \leftarrow w + a/w$
- DROP: $w \leftarrow w - b*w$
- SLOW-START: $w \leftarrow w + c$
- $a = 1; b = 0.5; c = 1$



HSTCP Algorithm

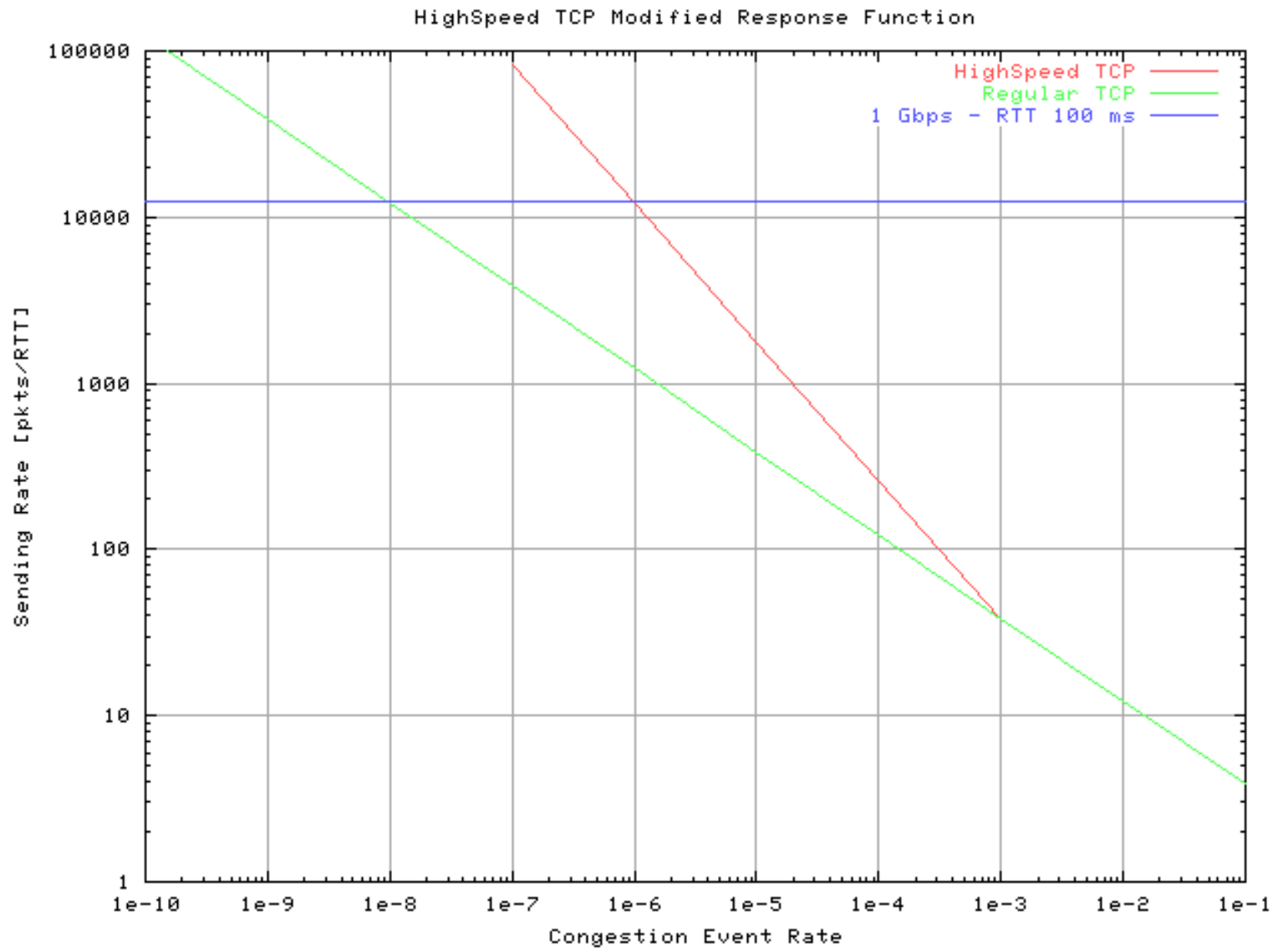
HSTCP algorithm:

$$w \leftarrow w + a(w)/w$$

$$w \leftarrow w - b(w) * w$$

$$a(w) = \frac{HWin^2 \times HP \times 2 \times b(w)}{2 - b(w)}$$

$$b(w) = \frac{(HDecr - 0.5) \times (\log(w) - \log(W))}{\log(W1) - \log(W)} + 0.5$$





Objective and Questions

Objective

Study the deployment of HSTCP in high speed long distance links as an effective mechanism for bulk data transfer, while maintaining fairness with other types of TCP already in use.

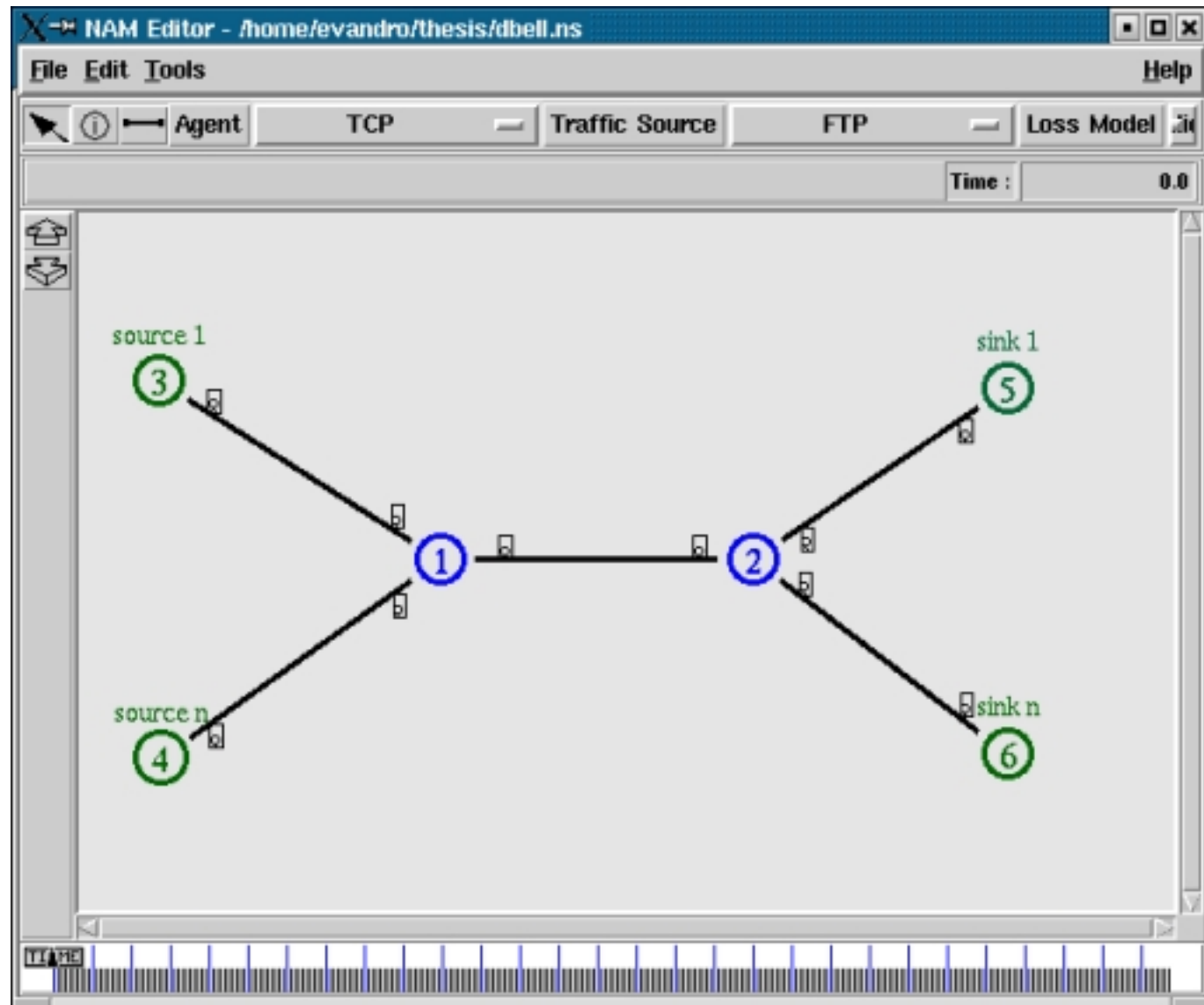
Main questions to this study:

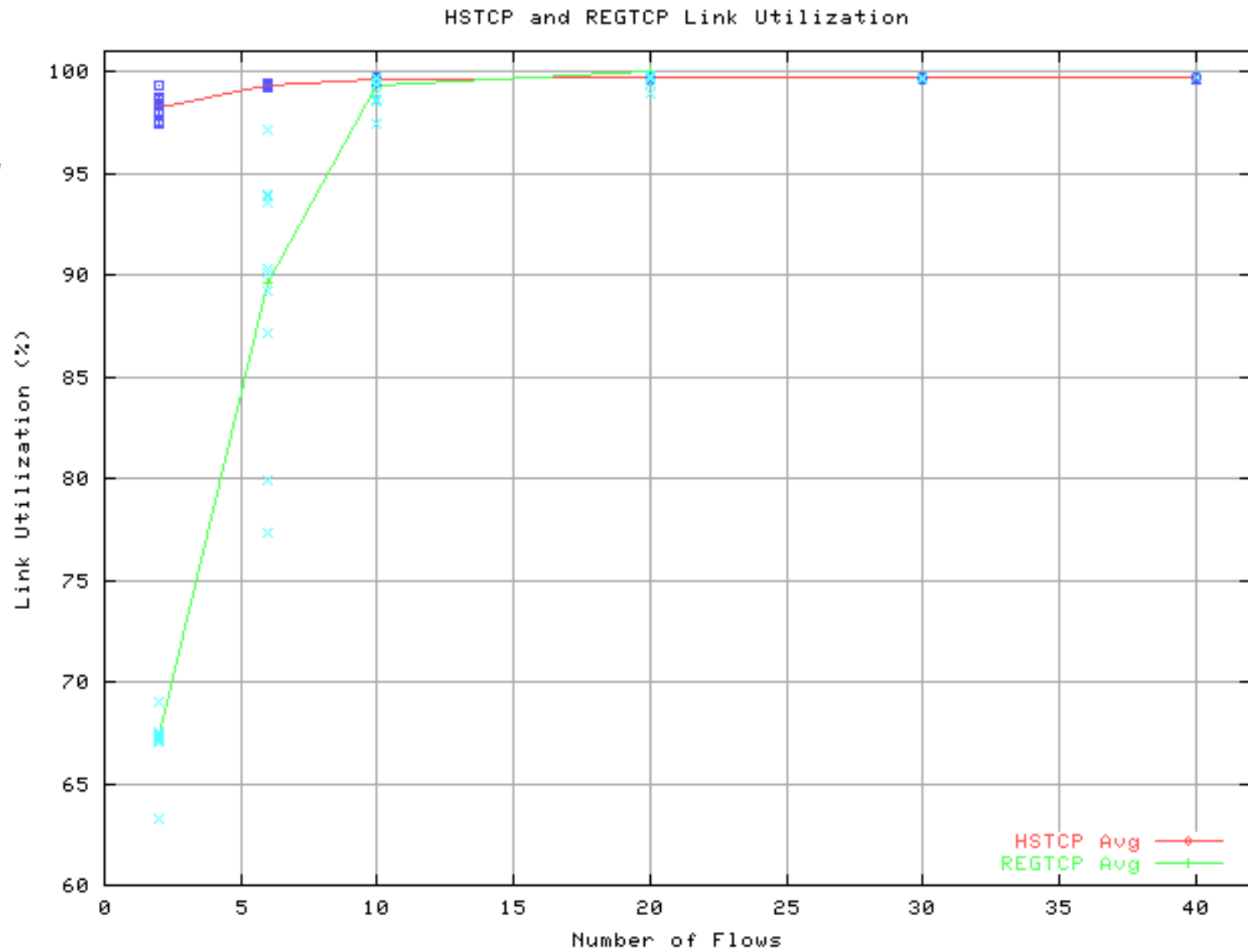
- A) What is the performance of HSTCP in situations where REGTCP has a weak performance?

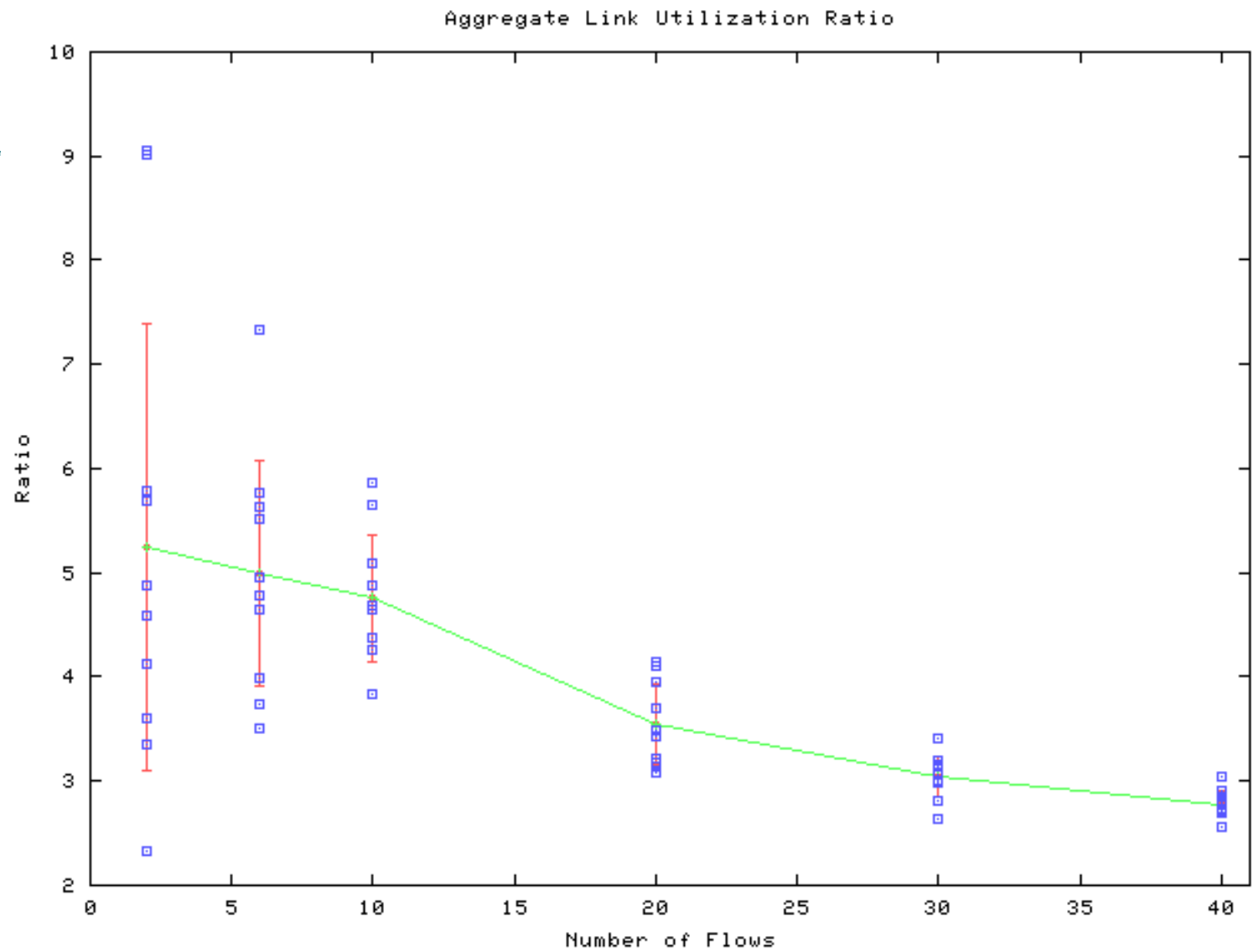


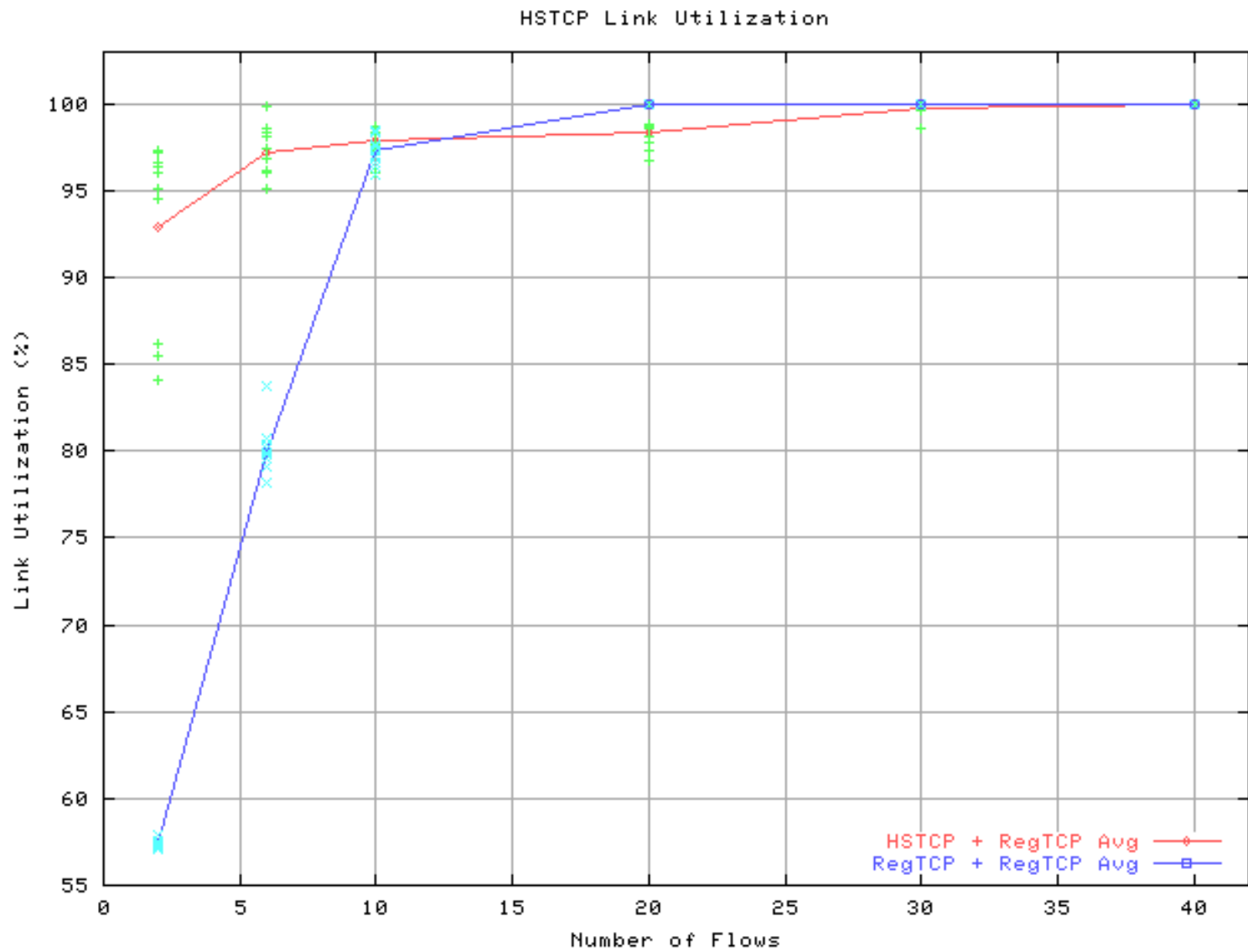
Main Questions

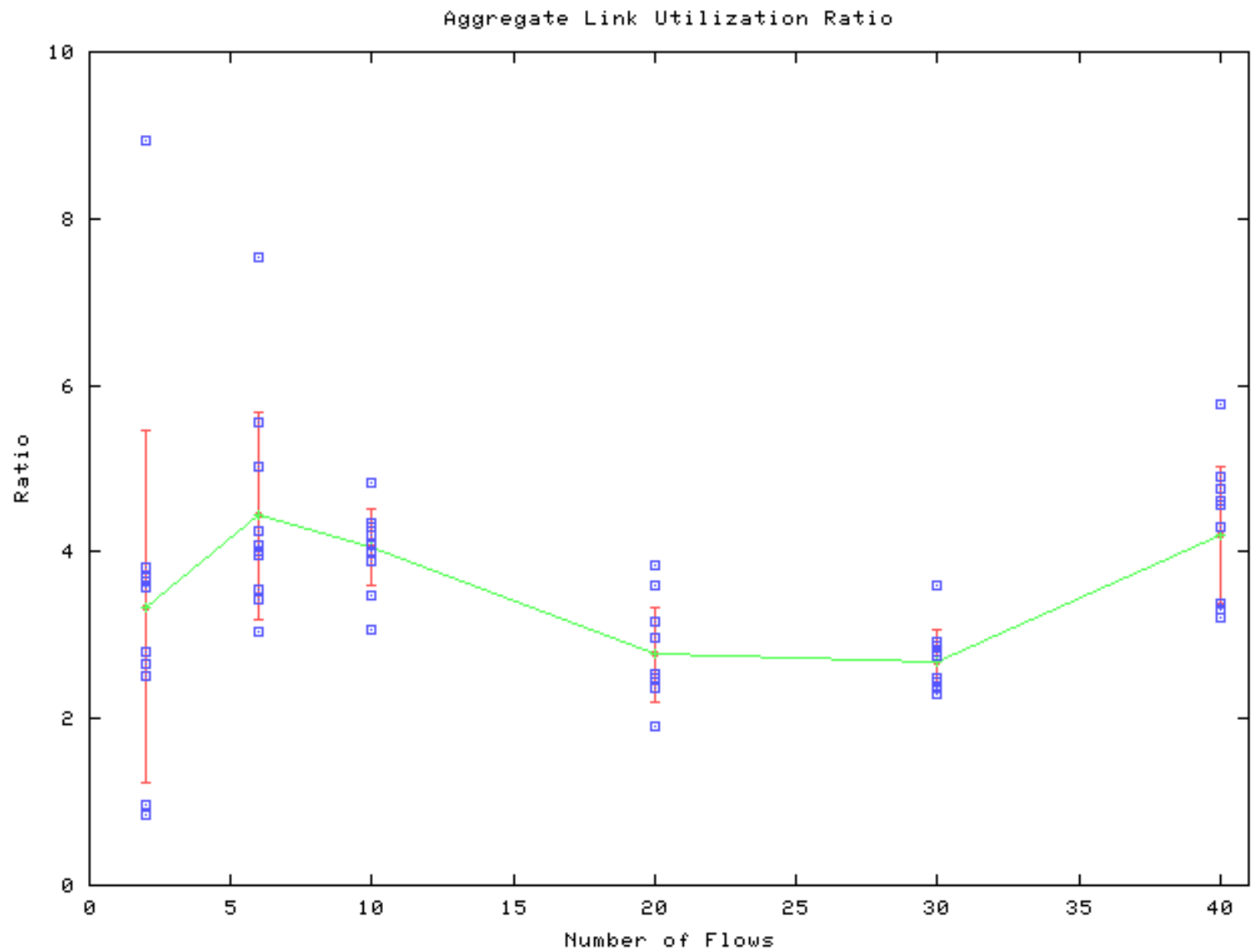
- B) Is possible to use HSTCP and REGTCP together and keep a relative fairness?
- C) What is the effect of the queuing police (RED/DT) in the performance of HSTCP and in the fairness between HSTCP and REGTCP?
- D) Can HSTCP be a substitute to other types of bulk data transfer?

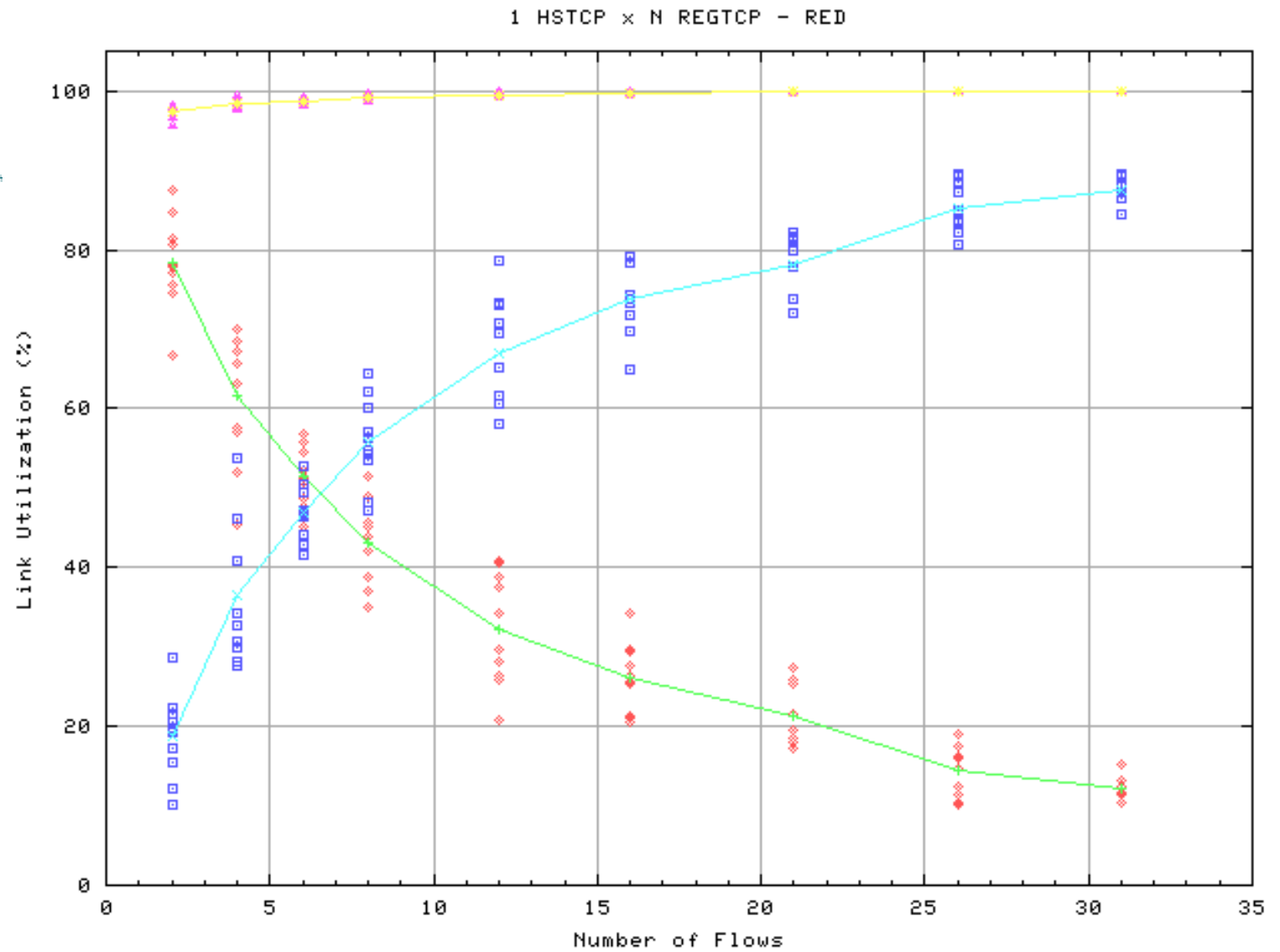














Present Investigations

Behavior with bursty traffic

Comparison with multiple TCP streams

Long-term HSTCP-REGTCP interaction



Conclusion

For the future DOE scientific networking is necessary a performance beyond what is available today

TCP presents severe restrictions for high bandwidth and long delay links

HSTCP is a proposal that presents interesting features and results

Is HighSpeed TCP a way?



References

A Vision for DOE Scientific Networking Driven by
High Impact Science - March 15, 2002

Simulation Study of HighSpeed TCP for Large
Congestion Windows - [http://www-
itg.lbl.gov/~evandro/hstcp](http://www-itg.lbl.gov/~evandro/hstcp)

HighSpeed TCP for Large Congestion Windows -
[http://www.ietf.org/internet-drafts/draft-floyd-tcp-
highspeed-00.txt](http://www.ietf.org/internet-drafts/draft-floyd-tcp-highspeed-00.txt) draft-floyd-tcp-highspeed-00.txt